A Bayesian Statistical Analysis of the DNA Contamination Scenario

John T. Wixted, Nicholas J.S. Christenfeld and Jeffrey N. Rouder*

ABSTRACT: When an unknown DNA sample recovered from crime-scene evidence is later found to match someone's known DNA profile, a question that sometimes arises is whether the match occurred because of a cross-contamination event in the lab. The concern is particularly acute if both samples were analyzed in the same lab at about the same time. Even if, upon review of the records, lab technicians can find no evidence of a mistake and are therefore completely certain that no mistake was made, the possibility that an undetected contamination event occurred cannot be logically dismissed. Intuitively, the difference between a zero chance of contamination and a very small chance of contamination (e.g., 1-in-1500) seems like the difference between certain evidence of guilt and *nearly* certain evidence of guilt. However, when DNA samples are contemporaneously analyzed in the same lab, the difference between a zero chance of contamination and a very small chance of contamination can instead be the difference between certain evidence of guilt and strong evidence of *innocence*. Here, we demonstrate that counterintuitive fact by applying a Bayesian statistical analysis to an unusual case where DNA contamination has long been suspected of having occurred despite testimony from lab technicians claiming that the probability of a contamination error was literally zero.

* John T. Wixted, Distinguished Professor of Psychology at the University of California, San Diego; Nicholas J.S. Christenfeld, Professor of Psychology at the University of California, San Diego; Jeffrey N. Rouder, Falmagne Endowed Chair of Mathematical Psychology and Professor of Cognitive Science at the University of California, Irvine.

When an unknown DNA profile found on crime-scene evidence is found to match someone's known DNA profile, a common concern is that it might be a false positive. A false positive occurs when the crime scene DNA was deposited by Person A but the profile is found to match Person B. However, if the DNA profile from the evidentiary sample is largely intact, the *random match probability* is such that the odds of a false positive are extremely remote (e.g., 1 in a trillion). Another concern that arises in some cases is not with that vanishingly unlikely possibility but instead with the possibility that an error occurred while the DNA was being analyzed, such as the mislabeling of evidence samples. For example, if a DNA profile obtained from a drinking glass is mislabeled as having come from a rape kit, the individual who drank from that glass will match and be suspected of having committed the rape. Using a Bayesian analysis, Thompson, Taroni and Aitken[1] showed that even a relatively low rate of human error in the lab (e.g., 1-in-10,000 tests) can introduce far more uncertainty about the value of DNA evidence than is implied by an extremely low random match probability. When there is concern that an error like that may have occurred in the lab, a National Research Council (NRC) report argued that the best approach is to retest the relevant samples.[2]

Other potential lab errors, such as cross-contamination, cannot be corrected so easily. As in the case of mislabeling, the problem would not be that the DNA does not belong to the matching individual (i.e., it would not be a false positive). Instead, it would definitely belong to the matching individual. However, appearances notwithstanding, that fact would not indicate that the matching individual was present at the crime scene. Moreover, as noted by Thompson et al.[3] (2003), re-testing will not solve the problem. If, for example, the evidentiary sample was

---

[1] William C. Thompson, F. Taroni. & C.G.G. Aitken, *How the probability of a false positive affects the value of DNA evidence*, 48 JOURNAL OF FORENSIC SCIENCES 47-54 (2003).
[2] National Research Council Committee on DNA Forensic Science: An Update, The evaluation of forensic DNA evidence, National Academy Press, Washington DC, (1996).
[3] *See Supra* note 1 at 2.

contaminated in the lab by the suspect's reference sample, then re-testing will only yield the same misleading result no matter how many times it is repeated.

In a case where there is concern about the possibility of laboratory contamination, the competing hypotheses differ from the usual hypotheses, which focus on whether the evidentiary DNA belongs to the matching individual (Hypothesis 1) or to someone else (Hypothesis 2). In the contamination scenario, the competing hypotheses are “the matching individual’s DNA was deposited on the evidence at the crime scene at the time the crime was committed” (Hypothesis 1) vs. “the matching individual’s DNA was deposited on the evidence due to a contamination event in the lab” (Hypothesis 2). Here, we provide a Bayesian analysis of the contamination scenario that compares those two hypotheses. To illustrate why such an analysis can be important, we use an actual case where lab contamination has long been suspected of having occurred. As we show later, the importance of the Bayesian analysis is that it demonstrates that if one assumes no chance of lab contamination (i.e., the contamination rate literally equals 0), then the outcome will support the first hypothesis. However, under certain conditions, if one allows for even a small chance of contamination (e.g., the contamination rate = 1-in-1500, meaning that 99.93% of analyses are error-free), then, counterintuitively, the outcome strongly supports the second hypothesis instead. The conditions under which that reversal occurs are rare, but they were present in the case of Gary Leiterman, where the possibility of contamination in the lab has

frequently been raised as a concern.[4] Indeed, this case is one of four DNA contamination cases highlighted in a National Institute of Justice report entitled "DNA for the Defense Bar."[5] We therefore use this case to illustrate the point that the difference between a contamination rate of literally 0 and a contamination rate as small as 1-in-1500 can be the difference between concluding that the matching individual was at the crime scene (and is therefore likely guilty) versus the matching individual is the victim of a contamination event (and is therefore likely innocent).

The Conviction of Gary Leiterman

Gary Leiterman, in 2005, was convicted of having murdered University of Michigan graduate student Jane Mixer in March of 1969 and was sentenced to life in prison without the possibility of parole. His conviction was based on a cold case analysis conducted in early 2002 (33 years after the crime) in the Michigan State Police Laboratory. An unknown DNA profile found on Jane Mixer's pantyhose was entered into a DNA database, where it matched the known profile of Gary Leiterman. His known profile had been entered into the database following his arrest for forging a prescription for pain medication in late 2001 at the age of 61. Following his arrest, he provided a reference DNA sample (saliva), which was then analyzed for inclusion in the Michigan DNA database. The match between Leiterman's known DNA profile and the unknown evidentiary DNA profile from Mixer's pantyhose was the main evidence leading to his

---

[4] Simon A. Cole & M. Lynch. *The Social and Legal Construction of Suspects*, 2 ANNUAL REVIEW OF LAW & SOCIAL SCIENCE 39–60 (2006); Chrsitopher Halkides & K. Lott, *Presumptive and Confirmatory Blood Testing*. In W. J. Koen & C. M. Bowers (Eds.), Forensic science reform: protecting the innocent (pp. 246-269). London: Academic Press (2017); Erin E. Murphy *Inside the cell: the dark side of forensic DNA*. New York, NY: Nation Books (2015); Andrea Roth, *Safety in Numbers - Deciding When DNA Alone is Enough to Convict*, 85 N.Y.U. L. REV. 1130 (2010); Andrea Roth, *Defying DNA: Rethinking the Role of the Jury in an Age of Scientific Proof of Innocence*, 93 B.U. L. REV. 1643 (2013); Boaz Sangero & M. Halpert, *Why a Conviction Should Not Be Based on a Single Piece of Evidence: A Proposal for Reform*, 48 JURIMETRICS J. 43–94 (2007); William C. Thompson. *Tarnish on the 'gold standard:' Understanding recent problems in forensic DNA testing*, 30 THE CHAMPION 10-16 (2006).

[5] Eric H. Holder, M. L. Leary & J. H. Laub, *DNA for the Defense* Bar. National Institute of Justice. https://www.ncjrs.gov/pdffiles1/nij/237975.pdf at 137 (2012).

conviction in 2005. Other evidence against him included the fact that he owned a .22 caliber gun in 1969 (Mixer was shot in the head with a .22), as well as the fact that a handwriting expert called by the prosecution testified that two words written on a phonebook near where Mixer was abducted were written by him (but a different handwriting expert called by the defense concluded that those words were not written by him).[6] By all accounts, the DNA evidence was decisive. The occurrence of that DNA match is what changed Leiterman's status from being just one of millions of people who might have committed the crime to being someone who is now serving a life sentence as Jane Mixer's murderer.

As noted above, serious reservations about the DNA evidence in this case have long been expressed.[7] The main reason for those reservations is that there were actually two unknown DNA profiles obtained from the Mixer crime-scene evidence analyzed in 2002. The second unknown profile was from a blood spot taken from Jane Mixer's left hand back in 1969. When that unknown profile was entered into the DNA database, it matched the known profile of John Ruelas. His profile had been entered into the database following his arrest for the murder of his mother in early 2002. The fact that he murdered his mother would ordinarily make him a strong suspect for the murder of Jane Mixer as well except that, in 1969, Ruelas was only 4 years old. His young age at the time of Mixer's murder obviously ruled him out as a suspect. Still, the notion that a 4-year-old preschooler was present at the after-midnight murder scene and bleeding on the victim struck many as being entirely implausible and raised the specter of DNA contamination in the lab (no connection between Leiterman and Ruelas was ever established

[6] A National Research Council committee found that there has been "…only limited research to quantify the reliability and replicability of the practices used by trained document examiners" and that "The scientific basis for handwriting comparisons needs to be strengthened," although the committee also acknowledged there may be some value to handwriting analysis (Strengthening Forensic Science in the United States: A Path Forward, National Academy Press, Washington DC, 2009).

[7] *See supra* notes 4 and 5.

despite an exhaustive inquiry). Reinforcing that concern was the troubling fact that DNA samples from all three individuals – Mixer, Leiterman, and Ruelas – were independently analyzed at about the same time (in early-to-mid 2002) in the Michigan State Police Laboratory. According to laboratory records, Mixer's cold-case evidence was analyzed in March and April of 2002; Leiterman's buccal swab arrived at the lab in February of 2002 and was analyzed in July of 2002; crime-scene evidence from the Ruelas murder arrived at the lab in February of 2002 and was analyzed in late March of that year. According to the prosecution's theory, Mixer, Leiterman and Ruelas were together on the night that Mixer was murdered in 1969 between midnight and 3 am (with 26-year-old Leiterman being the murderer and 4-year-old Ruelas being a bleeding bystander) and, due to an extraordinary coincidence arising from chance alone, the DNA samples from all three were together again in early 2002 in the Michigan State Police Laboratory. This theory further holds that no contamination occurred, a claim that was emphatically endorsed by lab technicians who testified that there was no possibility of a cross-contamination event in this case. According to the defense's theory, by contrast, DNA from Leiterman and Ruelas was inadvertently deposited on the Mixer cold case evidence in 2002 due to an undetected contamination event. The jury agreed with the prosecution's theory in 2005, and an appeals court ruled in 2007 that the trial was properly conducted. Leiterman is currently serving a life sentence without the possibility of parole.

Bayesian Analysis

Bayesian inference is a widely accepted statistical method of inductive reasoning based on the reassessment of competing hypotheses in light of new evidence.[8] The statistical analysis we present compares two competing hypotheses: the prosecution's hypothesis that Leiterman's

[8] Bayesian inference, https://www.nature.com/subjects/bayesian-inference

DNA was deposited on Mixer's clothing at the crime scene in 1969 versus the defense's hypothesis that Leiterman's DNA was deposited on Mixer's clothing in the DNA lab in 2002. Each hypothesis involves several inputs that bear on the likelihood that the hypothesis is true, and those inputs determine the outcome of the analysis (pointing in favor of guilt or innocence). All of the inputs in our analysis are explicit and have an empirical basis. Most are intuitively sensible to us, but this does not mean that they are necessarily the best estimates. However, having them explicit makes them easy to challenge (and change) if any seem unreasonable or unduly biased in favor of the prosecution or the defense.

It is important to emphasize at the outset that our main analysis is based on the DNA evidence, not on any other evidence that is potentially relevant to this case. For example, this new statistical analysis does not take into consideration the fact that Leiterman owned a .22 caliber gun in 1969 (which is unrelated to the DNA analysis and points in the direction of Leiterman's guilt) or the fact that a serial killer who was operating in the area at the time (John Norman Collins) killed another University of Michigan graduate student 10 weeks after the Mixer murder, by shooting her in the head with a .22 caliber gun (which is also unrelated to the DNA analysis and points in the direction of Leiterman's innocence). Instead, the main analysis presented here ignores all other non-DNA evidence and makes the assumption that a determination of guilt or innocence hinges on the strength of the DNA evidence.[9] We focus primarily on the DNA evidence to underscore the counterintuitive point that the difference between a contamination rate of 0 and a contamination rate as small as 1-in-1500 can be the difference between concluding that the matching individual is guilty or innocent.

[9] The only other slightly incriminating evidence presented at the trial consisted of a decades-old memory of a former roommate according to which Leiterman collected newspapers containing articles about a serial killer suspected of having committed several other murders in the Ann Arbor area during the late 1960s.

A key input to our statistical analysis is the probability that a DNA analysis would result in a cross-contamination event. Such events are undoubtedly rare, but they do occur. According to a National Research Council report on DNA technology, “Laboratory errors happen, even in the best laboratories and even when the analyst is certain that every precaution against error was taken”.[10] Similarly, Gill and Kirkham argued that "...it should be recognized that laboratory contamination is impossible to avoid completely but its extent is generally unknown unless proactively assessed—the probability of contamination must always be greater than zero".[11] Yet, at the 2005 Leiterman trial, lab personnel testified that cross contamination was impossible because every precaution had been taken to guard against contamination errors. In terms of our analysis, if the lab technicians were right, it would mean that the estimated contamination rate for the Leiterman case should be set to 0. If that were true, then the only reasonable conclusion would be that the Leiterman’s DNA was deposited on Mixer’s clothing at the crime scene (consistent with the prosecution's theory).

The statistical analysis we present below returns a probability of 1.0 that the prosecution’s theory is correct when the contamination rate is assumed to be 0 (as it must). However, our analysis also shows that changing the contamination rate from 0 (i.e., 100% error-free) to a value as low as 1-in-1500 (i.e., 99.93% error-free) results in an outcome that completely reverses the conclusion. That is, if the contamination rate is only 1-in-1500, such that 99.93% of analyses are error-free, then the estimated probability is close to 1.0 that the *defense’s* theory is true. Intuitively, by contrast, the difference between 100% of DNA analyses being free of cross-contamination errors vs. 99.93% of DNA analyses being free of cross-contamination

---

[10] National Research Council, *DNA technology in forensic science*, National Academy Press, Washington DC, (1992) at 89.

[11] Peter Gill & A. Kirkham, *Development of a Simulation Model to Assess the Impact of Contamination in Casework Using STRs*. 49 JOURNAL OF FORENSIC SCIENCE 485-491 (2004).

errors seems like the difference between certain guilt vs. *almost* certain guilt. Contrary to that powerful intuition, under some circumstances – such as the circumstances that apply to the Gary Leiterman case – it can be the difference between certain guilt vs. almost certain innocence (where, in this case, "guilt" refers to having deposited DNA at the crime scene, and "innocence" refers to the DNA having been deposited by accident in the crime lab). The formal analysis is presented next.

The Competing Hypotheses and Relevant Data

We consider the following three mutually exclusive and exhaustive hypotheses:

- H_1: Leiterman's DNA was deposited on Mixer's clothing only at the crime scene in 1969
- H_2: Leiterman's DNA was deposited on Mixer's clothing only in the lab in 2002
- H_3: Leiterman's DNA was deposited on Mixer's clothing at the crime scene in 1969 and in the lab in 2002

The probability of H_3 is obviously extremely low because it assumes that the person who likely murdered Jane Mixer in 1969 just happened to be the same person whose DNA contaminated her evidence in the lab 33 years later, but we include it for completeness. The bulk of our analysis focuses on the odds of H_1 vs. H_2, the two plausible hypotheses, but we formally consider H_3 when we ultimately compute the estimated probability of guilt even though its likelihood is much too low to appreciably affect the end result.

The next step is delineating the relevant data for analysis. According to our reading of the trial transcripts, there were four main observations, what are called "events" (E) that comprise the relevant data. The following four events are not disputed by either the defense or the prosecution:

- E_m: There was a definite match (hence, the subscript m) between Leiterman's known

profile and the DNA found on Mixer's pantyhose.

- E_s: The matching DNA on Mixer's pantyhose was consistent with saliva (*s*). That is, trial testimony from lab technicians indicated that Leiterman's DNA on Mixer's pantyhose was not consistent with blood or semen but was consistent with other biological materials, including saliva.
- E_e: Leiterman's DNA was exclusively (*e*) found on Mixer's pantyhose (i.e., there was no DNA from Mixer herself).
- E_c: Leiterman's reference sample following his arrest in 2001 (a saliva sample) and Mixer's cold-case crime evidence from 1969 were in the lab contemporaneously (*c*) in the first 6 months of 2002, though they were not analyzed on the same days.

A key component of our analysis consists of estimating the probability of each event assuming H_1 and, separately, assuming H_2. We make the assumption that the four events are statistically independent. Therefore, whether $H = H_1$ or $H = H_2$:

$$P(D|H) = P(E_m|H) \times P(E_s|H) \times P(E_e|H) \times P(E_c|H) ,$$

where *P* indicates the probability of the event in parentheses, and the vertical bar indicates a conditional probability. Thus, for example, $P(D|H_1)$ represents the probability of the data given Hypothesis 1.

By the Law of Conditional Probability:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)} \text{ x } \frac{P(H_1)}{P(H_2)}$$

This equation is Bayes' rule in odds form. The ratio on the left-hand side, the relative probability that the hypotheses are true in light of the relevant data, is the target of interest. It is

called the “posterior odds.” The term on the far right is the prior odds, the probability that the hypotheses are true before any data are observed. The middle term is the likelihood ratio, and it describes how the data have influenced the probabilities that the hypotheses are true. It is helpful here to restate Bayes rule in verbal form:

$$\text{posterior odds} = \text{likelihood ratio} \times \text{prior odds}$$

The posterior odds of H_1 vs. H_2 is the target of interest for assessing the guilt of Leiterman. For example, if the outcome of the analysis yields a posterior odds of 100 / 1, it would mean that the odds are 100-to-1 in favor of H_1 (the prosecution's hypothesis) given the data. To compute the posterior odds of guilt (which can then be translated into a posterior probability of guilt), we need to first compute both the prior odds and the likelihood ratio.

The Prior Odds

The prior odds are equal to $P(H_1)/P(H_2)$. $P(H_1)$ is the probability that, before anything is known about the identity of the person whose DNA was deposited on Mixer’s clothing in 1969, it would turn out to be Leiterman. The prior probability that the matching individual would turn out to be Leiterman is set to $P(H_1) = 1/N$, where N is the number of people who, had their DNA matched the unknown DNA profile on the Mixer crime scene evidence and before anything else was known about their personal history (e.g., about potential alibis), would have been suspected of having deposited their DNA at the crime scene.

To estimate N, the first step is to set the radius around Ann Arbor, Michigan, (where Mixer lived) within which plausible candidates for having been at the crime scene resided. We set this radius based on the prosecution’s interpretation of the DNA evidence. Specifically, both Leiterman and Ruelas were judged by the prosecution to have deposited their DNA at the crime

scene. Ruelas lived in Detroit in 1969 about 40 miles from Ann Arbor, and Leiterman lived in the outskirts of Detroit about 20 miles from Ann Arbor. We therefore set a radius of 40 miles around Ann Arbor, which just barely includes where both Leiterman and Ruelas lived in 1969. At a minimum, anyone living in that region in 1969 whose DNA was found on Mixer's evidence would have been judged by the prosecution to have been at the crime scene. It seems reasonable to suppose that the radius (and the corresponding population) would actually be considerably larger than that, but we use a 40-mile radius as a conservative estimate. It is a conservative estimate (favoring the prosecution) because if the DNA had matched an adult male who lived as far away as 80 miles from the crime scene in 1969, it seems likely he would have been regarded as a viable suspect.

Today, the population living within a 40-mile radius of Ann Arbor is approximately 4.4 million people. According to the US Census Bureau, the entire population of Michigan in 1969 was ~8.8 million people, which is about 89% of what it is today. Thus, a reasonable value for this parameter would be $N = .89 \times 4.4$ million people ≈ 4 million people. Keep in mind that this is an estimate of the number of people who, before the fact, would have been judged by the prosecution to have deposited their DNA at the crime scene had their DNA been found on Mixer's crime scene evidence in 2002 (as was true of both Leiterman and Ruelas).

The ultimate question is whether or not Leiterman murdered Mixer. Thus, if we assume that H_1 is that Leiterman deposited his DNA at the crime scene *because he murdered her*, then many of these 4 million people would be excluded for the same reason that Ruelas was excluded as being the murderer (even though he was not excluded as having deposited his DNA at the crime scene). To be considered as a plausible suspect for the murder, it seems likely that the individual whose DNA was found on the Mixer evidence would have had to have been an adult

male living in the region at the time of the murder. Approximately 50% the Michigan population is male, and approximately 50% fall between the ages of 18 and 59. Thus, we can reasonably set $N = .50 \times .50 \times 4$ million = 1 million. This is an estimate of the number of people who, before anything more than age and sex was known about them, would have been judged by the prosecution to have deposited their DNA at the crime scene had their DNA been found on Mixer's crime scene evidence in 2002 and who would have been suspected of having murdered her (as Leiterman was, but not Ruelas was not). That is, in the absence of any further individuating evidence, each of those adult men had about a one in a million chance of being the killer.

A disadvantage of this approach to estimating N is that it could be made arbitrarily large or small to suit one's ends. For example, to minimize N, favoring the prosecution, one could more narrowly define it to be the number of males between the ages of 17-44 in 1969 who lived within 20 miles of the crime (as Leiterman did) and who were not married (as Leiterman was not). However, in the appendix, we describe an alternative approach to estimating N that relies solely on DNA database search statistics (not on any defense- or prosecution-oriented assumptions) and that arrives at approximately the same value we arrived at above (~1 million). In addition, although we use N = 1 million, we later show that our conclusions hold even when much lower values of N are assumed (i.e., even when the value is implausibly biased to favor the prosecution's theory). For our main analysis, $P(H_1) = 1 / 1 \text{ million} = 10^{-6}$.

We turn next to estimating the prior probability of the defense's hypothesis, $P(H_2)$. This is the probability that, before anything is known about who matched, Leiterman would be the one whose DNA was deposited on Mixer's clothing. What is the probability that, given a

contamination event, the person involved would be Gary Leiterman? According to Jen[12], 9933 reference samples were analyzed in Michigan lab in 2002. We therefore assume that approximately half that number ($n \approx 5000$) were analyzed contemporaneously with the Mixer evidence in the first half of 2002. Thus, the probability that Leiterman is the one of these 5000 is $P(H_2) = 1 / n = 1 / 5000 = 2 \times 10^{-4}$. Note that this number could be reasonably set to a smaller value because we have not reduced it by limiting it to males of the appropriate age. However, reducing it would favor H_1 (the defense's theory), so we conservatively assume that all 5000 would be plausible suspects even though the age (too young) and sex (female) of at least some would exclude them.

Having estimated $P(H_1)$ and $P(H_2)$, we are in a position to estimate the prior odds, $P(H_1)/P(H_2)$. With $P(H_1) = 10^{-6}$ and $P(H_2) = 2 \times 10^{-4}$, the prior odds come to $P(H_1)/P(H_2) = 0.005$ (i.e., 1-in-200). This value means that, before knowing anything else about the evidence considered below, a DNA match would imply that if Leiterman matched, it was 200 times more likely that his DNA was deposited in the lab due to a contamination event than at the crime scene while murdering Jane Mixer. Hence, the prosecution has a large burden. But if the probability of contamination is 0, the DNA match overcomes this burden. The question of interest is whether the same is true when the probability of contamination is close to 0 (e.g., 1-in-1500) but not exactly 0.

The Likelihood Ratio

We next turn to estimating the parameters needed to compute the likelihood ratio. To do so, we compute the probability of observing each of the four relevant events given H_1 and, separately, given H_2.

[12] See supra note 12.

***E_m* (the DNA Match)**. Under what conditions would a match occur assuming H_1? This is the probability that Mixer's murderer would deposit his DNA on her clothing in the first place (p_1) times the probability that the DNA would survive 33 years of storage (p_2). That is, $P(E_m|H1) = p_1 \times p_2$. We first estimate p_1. According to Jen[13], in the Michigan State Police Laboratory circa 2002, only half of the recent (i.e., not cold case) forensic evidence received by the lab yielded a DNA profile that was sufficiently intact to enter into a database. This means that only half the time, when a crime is committed, does someone other than the victim (presumably the perpetrator in most cases) leave an identifiable DNA profile on the evidence. A similar result was observed in a prospective study of DNA collected from burglary crime scenes.[14] Out of 1079 cases, they found that 54.7% resulted in the generation of a profile that was sufficiently intact to search a database. Thus, we assume that $p_1 = 0.50$.

Consider next the probability that DNA on crime scene evidence would survive 33 years of storage in the first place (p_2). A news story published in 2011 about a cold case unit in Iowa provides some guidance on this issue. The story indicated that "…the unit collected data on 150 unsolved cases going back to the 1960s. The investigators prioritized 50 of them and conducted DNA analysis on 2,018 pieces of evidence. They developed profiles in 23 cases and checked 11 of those profiles against the nationwide DNA database." In other words, 11 of the 50 cold cases subjected to DNA analysis yielded a DNA profile sufficiently intact to enter into a database (11/50 = .22), as happened in the Mixer cold case investigation. For reasons described above, we would expect only half the cases to yield a usable DNA profile even if it were a recent case involving little to no degradation (because DNA is found on recent evidence only 50% of the

[13] Kyle I. Jen, *Michigan's forensic DNA database*. www.house.mi.gov/hfa/Archives/PDF/dna.pdf (2003) at 5.
[14] John K. Roman, S. Reid, J., Reid, A. Chalfin, W. Adams, & C. Knight, *The DNA field experiment: a randomised experiment of the cost-effectiveness of using DNA to solve property crimes*. 5 JOURNAL OF EXPERIMENTAL CRIMINOLOGY 345-369 (2008).

time). It therefore follows that, for the ~50% of cases in which DNA was initially deposited on the evidence, it degraded in storage about half the time (i.e., .50 ×.50 = .25, which is close to the observed value of .22). Thus, we assume DNA would survive cold-case long-term storage with a probability of .50 (p_2 = .50), which seems like an intuitively sensible value. If anything, it is probably conservative (favoring an outcome of guilt rather than innocence) since the cases investigated by the Iowa cold case unit were probably less than 33 years old, on average. For our analysis, we set the probability of DNA being deposited by the perpetrator while committing the crime and surviving 33 years of storage assuming H_1 to $P(E_m|H_1) = p_1 \times p_2 = 0.25$.

The probability of a match given H_2, $P(E_m|H_2)$ is the probability that Leiterman's DNA was deposited on Mixer's clothing (and would therefore match) due to a contamination event in the lab. The probability of a contamination event is initially set to 1 / 1500 in our main analysis based on recent research by Kloosterman et al.[15] They reviewed lab records for 472,127 DNA analyses conducted at the Human Biological Traces Department of the Netherlands Forensic Institute (NFI) over the years 2008–2012 (see their Table 1). Their Table 5 shows that cross-contamination events with other samples in the lab occurred 311 times in the 2008-2012 time period, which is to say that such an event occurred in 311 / 472,127 ≈ 1 / 1500 analyses. Note that these are *detected* cross-contamination events, most of which were caught by the NFI quality control systems. Our actual interest is in estimating the rate of *undetected* cross-contamination events – the kind of event that, according to H_2 (the defense's contamination theory), occurred in the Leiterman case. There is no easy way to precisely determine how often undetected events occur because it would require blind testing of lab technicians, which, given the rarity of

[15] Ate Kloosterman, M. Sjerps & A. Quak, *Error rates in forensic DNA analysis: Definition, numbers, impact and communication*. 12 FORENSIC SCIENCE INTERNATIONAL: GENETICS 77-85 (2014).

contamination events, would be a prohibitively expensive undertaking.[16] Nevertheless, the results reported by Kloosterman et al.[17] indicate that cross-contamination events do occur at a low rate. This finding underscores a point stressed in the 1996 NRC report:[18] "No amount of effort and improved technology can reduce the error rate to zero." It seems reasonable to suppose that the error rate reported by Kloosterman et al.[19] gives us a ballpark estimate of the value we are interested in. Based on the assumption that the rates of detected and undetected contamination rates are similar, we initially set $P(E_m|H_2) = 1 / 1500$. Later, we consider how the results change when its value set to within an order of magnitude of that starting value (i.e., 1/150 and 1/15,000).

***E_s* (the DNA is consistent with saliva)**. The second relevant event, E_s, is that the DNA on Mixer's pantyhose was not consistent with blood or semen but was consistent with saliva. We first consider $P(E_s|H_1)$, which is the probability that E_s would be observed given H_1 (the prosecution's theory). If Leiterman killed Mixer (H_1), the DNA on her pantyhose did not have to be consistent with saliva. Had the DNA been from semen, for example, a match to his known reference sample (from saliva) would still have occurred. Thus, according to H_1, consistency with saliva was a coincidence that occurred by chance with some probability. We estimate the probability of such a coincidence using data reported by Cross et al.[20]

Cross et al. examined forensic evidence from a large number of sexual assault cases in Massachusetts. The Mixer murder did not involve rape but did appear to have a sexual

---

[16] *See supra* note 1 at 7.
[17] *See supra* note 14.
[18] *See supra* note 2 at 4.
[19] *See supra* note 14.
[20] Theodore P. Cross, M. Alderden, A. Wagner, L. Sampson, B. Peters, M. Spencer & K. Lounsbury, K., *Forensic evidence and criminal justice outcomes in a statewide sample of sexual assault cases.* https://www.ncjrs.gov/pdffiles1/nij/grants/248254.pdf (2014).

motivation given that her pantyhose had been pulled down.[21] Their Table 5.15 (p. 91) shows the percentage of cases in which various kinds of biological evidence were detected. According to that table, blood was detected 24.1% of the time, saliva 35.7% of the time, semen 62.5% of the time, and other biological materials 44.4% of the time. These percentages add up to more than 100% because a single case can yield more than one type of biological evidence. From these data, one can reasonably infer that approximately 50% of the time, when only one kind of biological evidence is found, that evidence would not be consistent with blood or semen but would be consistent with saliva and other biological materials (as was true of the Leiterman case). This is not an exact estimate given that the biological sources may be non-independent, but it seems like a reasonable estimate, one that accords with intuition. Thus, we set $P(E_s|H_1) = 0.50$.

Next consider $P(E_s| H_1)$, which is the probability that E_s (the forensic DNA was consistent with saliva) would be observed given H_2 (the defense's theory). H_2 holds that saliva from Leiterman's buccal swab is what contaminated the Mixer evidence. In other words, according to H_2, the probability is 1.0 that the DNA found on Mixer's pantyhose would be consistent with saliva. Thus, $P(E_s| H_2) = 1.0$. This setting ensures that had the DNA come from a source not consistent with saliva (such as blood or semen), the defense's contamination theory would be conclusively ruled out.

***E_e* (Leiterman's DNA, but not Mixer's DNA, found on Mixer's Clothing)**. The third relevant event, E_e, is the observation that DNA on the pantyhose came exclusively from Leiterman, with no detectable DNA from Mixer herself. Intuition suggests that the opposite

[21] Earl James, *Catching serial killers: Learning from past serial murder investigations*. International Forensic Services, Inc.: Lansing, Michigan. (1991).

would usually be true (i.e., that there would be more of Mixer's DNA on her own pantyhose than Leiterman's DNA).

Assuming the pantyhose were stored in such a way that the DNA would survive for 33 years (which it was according to H_1 given that Leiterman's DNA survived), what is the probability that more of Leiterman's DNA was deposited on Mixer's pantyhose than from Mixer herself? Two recent research articles offer some guidance about the probability of finding the outcome observed in the Leiterman/Mixer case, namely, a measurable amount of DNA from Leiterman (the "toucher") on Mixer's pantyhose but with no measurable DNA from Mixer (the "wearer"). Breathnach et al.[22] investigated "…the frequency of detection of DNA from wearer, toucher or others when individuals wore and handled worn garments under normal circumstances" (p. 53), and they reported that "Toucher and no wearer was observed in 15% of reportable samples" (p. 59). This suggests that a reasonable setting for $P(E_e|H_1)$ might be .15. Similarly, van den Berge et al.[23] investigated an activity scenario in which the toucher grabbed the trouser leg ankles of the wearer and dragged that individual for 1 minute. In 48 such cases, their Figure 4C indicates that an outcome somewhat similar to that observed in the Leiterman/Mixer case (that is, DNA from the toucher coupled with virtually no DNA from the wearer) occurred in approximately 7 cases. More specifically, in 10 of the 48 analyses, the toucher (called the "grabber" in that study) was the major contributor, and of those 10, 7 exhibited very small amounts of DNA from the wearer. In other words, once again, that outcome

[22] Michelle Breathnach, L. Williams, L. McKenna & E. Moore, *Probability of detection of DNA deposited by habitual wearer and/or the second individual who touched the garment,* 20 FORENSIC SCIENCE INTERNATIONAL: GENETICS 53–60 (2016).
[23] M. van den Berge, G. Ozcanhan, S. Zijlstra, A. Lindenbergh & T. Sijen, Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios. 21 FORENSIC SCIENCE INTERNATIONAL: GENETICS 81-89 (2016).

was observed with a probability of 7/48 = .15. We therefore we set $P(E_e|H_1) = .15$. Again, this is not an exact value, but it seems reasonable and accords with intuition.

According to H_2 (the defense's hypothesis), how likely is it that Leiterman's DNA would be detected on Mixer's pantyhose but her own DNA would not be detected? H_2 implies that Leiterman's DNA was recently deposited from his buccal swab and was detected for that reason. Thus, under this hypothesis, $P(E_e|H_2)$ reduces to the probability that Mixer's DNA would not be detected at all, as it was not.

Unlike H_1, H_2 does not imply that DNA deposited at the crime scene survived 33 years of storage. Thus, $P(E_e|H_2)$ is equal to the probability that Mixer's DNA was not deposited on her evidence in the first place (p_3) plus the probability that if it were deposited in the first place ($1 - p_3$) that it would degrade to the point of being undetectable after 33 years of storage (p_4). That is, $P(E_e|H_2) = p_3 + (1 - p_3) \times p_4$. From the toucher-wearer studies discussed above, we estimate that $p_3 \approx 0.15$ (i.e., wearers sometimes do not leave a trace of DNA on their own clothing). Earlier, we estimated that the perpetrator's DNA would survive 33 years of storage with a probability of $p_2 = 0.50$. Assuming the same is true for a victim's DNA on her own clothing, $p_4 = 0.50$. Thus, $P(E_e|H_2) = 0.15 + (1 - 0.15) \times 0.50 = 0.575$.

***E_c* (Contemporaneous DNA Analyses)**. The fourth relevant event, E_c, is that Leiterman's DNA happened to be in the lab during the same 6-month period in 2002 that the Mixer evidence was also in the lab and being analyzed. According to H_1 (the prosecution's theory), the contemporaneous analysis of the Leiterman and Mixer evidence is simply a coincidence that occurred with some probability less than 1.0 because Leiterman's buccal swab could have been analyzed at any time and a match to the DNA on Mixer's pantyhose would still

have occurred. According to Jen's report,[24] approximately n = 42,000 DNA analyses had been performed by the Michigan State Police Laboratory and entered into the database by the end of 2003. Leiterman's known profile was added some time in 2004, at which point the database likely had even more entries, but we conservatively assume n = 42,000 (if anything, favoring an outcome of guilt). Of all of the profiles in the database at the time Leiterman matched in 2004, what is the probability that the matching profile would have been contemporaneously analyzed with the Mixer evidence in the first 6 months of 2002? Under H_1, this outcome would occur by chance with a probability of 5000 / 42,000 = 0.12. Thus, we set $P(E_c|H_1) = 0.12$.

Note that this is a conservative value for another reason as well. The reason is that just as H_1 does not require that the Leiterman and Mixer analyses occur at the same time (which happened coincidentally with probability 0.12), it also does not require that the analyses happen in same *place*. For example, imagine that Leiterman murdered Mixer in 1969 (in accordance with H_1) and then moved to California shortly thereafter, where, at the age of 61 in 2001, he forged a prescription, was arrested, and had his known DNA profile entered in the California CODIS database. When the unknown DNA profile from Mixer's pantyhose was taken from her pantyhose in the Michigan State Police Crime lab in 2002 and then entered into the Michigan DNA database, it would not have yielded a match. However, the random match probability for Caucasians with this unknown profile was 1 in 171 trillion. That value far exceeds the criterion for entering the profile into the national (federal) CODIS database, which requires a random match probability at least as low as 1 in 10 million. Following a failed search of the Michigan CODIS database, had the unknown profile from the Mixer evidence been entered into the federal database, it would have matched the known profile of Gary Leiterman. He would then have been

[24] *See supra* note 12.

discovered to have lived in Michigan in 1969, and he would become a very strong suspect. And rightly so. Under these conditions, the possibility of contamination would be negligible (because, under this scenario, his buccal swab would have been analyzed in a California lab in early 2002, whereas the Mixer evidence was analyzed in a Michigan lab in early 2002). The point is just that H_1 (the prosecution's theory) does not require that the analyses were performed at either the same time *or* the same place. The coincidence factor of 0.12 used in our analysis only takes into account coincidental timing, not the additional fact of coincidental location (which would justify the use of an even lower value). Thus, the value of $P(E4|H1) = 0.12$ used in or analysis is, if anything, biased in favor of an outcome of guilt.

Under H_2, contemporaneous analysis in the same lab is a requirement (i.e., it had to happen for contamination to have happened). Thus, $P(E_c|\ H_2) = 1$.

Putting all of these estimates together, the numerator of the likelihood ratio is $P(D|H_1) = P(E_m|H_1) \times P(E_s|H_1) \times P(E_e|H_1) \times P(E_c|H_1) = 0.25 \times 0.50 \times 0.15 \times 0.12 = 0.00225$. Likewise, the denominator of the likelihood ratio is $P(D|H_2) = P(E_m|H_2) \times P(E_s|H_2) \times P(E_e|H_2) \times P(E_c|H_2) = .00067 \times 1.0 \times .575 \times 1.0 = .000383$. That is, $P(D|H_1) = 0.00225$ and $P(D|H_2) = 0.000383$. Thus, the likelihood ratio is equal to 0.00225 / 0.000383 = 5.87. This means it is approximately 6 times more likely that a DNA match would occur if H_1 were true compared to if H_2 were true.

The Posterior Odds and Posterior Probability

Given the estimates above of the prior odds and the likelihood ratio, we are now in a position to compute the posterior odds. Our point is not that the estimates provided above are indisputable. However, we assume that most would agree that they are at least reasonable and therefore allow us to reasonably estimate how the posterior odds change under different assumptions about the rate of contamination in the Michigan State Police Laboratory in 2002.

First, what are the posterior odds if we assume a contamination rate of only 1-in-1500 (i.e., 99.93% of DNA analyses are error free)? Using this estimated contamination rate, as we did in the analysis presented above, the likelihood ratio comes to 5.87. When multiplied by the prior odds of .005, the estimated posterior odds of H_1 vs. H_2 come to $P(H_1|D) / P(H_2|D) = .0294$. In other words, despite the fact that Leiterman's DNA matched the DNA found on Mixer's clothing, the odds of guilt are very low (which is to say that the odds of innocence are very high). Table 1 provides a summary of our Bayesian analysis.

To convert the posterior odds of guilt into a posterior *probability* of guilt, we need to consider the mutually exclusive and exhaustive possibilities H_1, H_2 and H_3, where H_3 assumes that Leiterman's DNA was deposited both at the crime scene in 1969 *and* in the lab in 2002. In truth, the probability of joint occurrence is so low that it will not affect our calculations in any appreciable way, but we mention it here for the sake of completeness.

Based on the DNA evidence, Leiterman would be guilty if either H_1 or H_3 were true. Thus, the posterior probability of guilt, $P(G|D)$, is equal to:

$$P(G|D) = [P(H_1|D) + P(H_3|D)] / [P(H_1|D) + P(H_2|D) + P(H_3|D)]$$

This equation contains three unknowns, $P(H_1|D)$, $P(H_2|D)$ and $P(H_3|D)$. Their values can be determined by algebraically solving the following three equations that relate these values to each other. From Table 1, we know that the posterior odds of guilt is given by $P(H_1|D)/ P(H_2|D) = 0.0294$, which means that

$$P(H_1|D) = 0.0294 \times P(H_2|D) \quad (1)$$

In addition, by definition,

$$P(H_3|D) = P(H_1|D) \times P(H_2|D) \quad (2)$$

Finally, because H_1, H_2, and H_3 are mutually exclusive and exhaustive, their values must sum to 1:

$$P(H_1|D) + P(H_2|D) + P(H_3|D) = 1 \quad (3)$$

Using these three equations to solve for the three unknowns yields the following estimates: $P(H_1|D) = 0.0285$, $P(H_2|D) = 0.9715$, and $P(H_3|D) = 2.19 \times 10^{-9}$. The value of $P(H_3|D)$ is so much smaller than the other two probabilities that the estimated value of $P(G|D)$ is virtually identical whether or not it is included. Thus, the posterior probability of guilt is closely approximated by:

$$P(G|D) = P(H_1|D) / [P(H_1|D) + P(H_2|D)]$$

This equation is simply a variant of the posterior odds that we computed earlier. As noted above, the posterior odds of guilt come to $P(H_1|D)/ P(H_2|D) = 0.02935$. Because probability = odds / (odds + 1), the posterior probability of guilt comes to 0.0294 / (1 + 0.0294) = 0.0285. If the posterior probability of guilt is 0.0285, the posterior probability of innocence is 1 - 0.0285 = 0.972.

Our analysis yields a completely different result if we now assume that the contamination rate is effectively 0 (not 1-in-1500). The lab technicians testified, in fact, that there was literally no chance of contamination in the Leiterman despite the disconcerting match to Ruelas. If the contamination rate, $P(E_m|H2)$, is set to a value approaching 0, then $P(D|H_2)$ would approach 0 as well. In that case, the likelihood ratio, $P(D|H_1) / P(D|H_2)$, and the posterior odds, $P(H_1|D) / P(H_2|D)$, would both approach infinity, which means that the posterior probability of guilt would approach 1.0. In other words, if we assume a contamination rate approaching 0 despite the Ruelas match, our statistical analysis would indicate that Gary Leiterman deposited his DNA on

Mixer's clothing in 1969 with probability approaching 1.0.

The critical point to take away from this analysis is that, contrary to a powerful intuition, it is not the case that 100% error-free DNA analyses implies that H_1 is certain, whereas 99.93% error-free implies that H_1 is *almost* certain. Instead, 100% error-free DNA analyses implies that H_1 is certain (as before), whereas 99.93% error-free implies that H_1 is almost certainly wrong and H_2 is almost certain.

Alternative Scenarios

Our central conclusion is reasonably robust to variations in some of the key estimates. For example, setting N (the number of plausible suspects prior to knowing who matched) to the implausibly low value of 250,000 – as if only 250,000 of the 8.8 million Michigan residents in 1969 would have become a suspect in the Jane Mixer murder had their DNA matched – only reduces the probability that H_2 is true to .90. Thus, for the analysis to yield any outcome close to favoring H_1, some other parameter would also have to be changed to an implausible value favoring that outcome.

Setting N back to a seemingly more reasonable value of 1 million, we can ask about the posterior odds with contamination rates within an order of magnitude of 1-in-1500. Setting the contamination rate to 1-in-150, the posterior probability of H_1 and H_2 come to 0.003 and 0.997, respectively. In other words, with a contamination rate that high, a decision based on the DNA evidence alone would indicate almost certain innocence. A contamination rate of 1-in-150 does not seem altogether unreasonable given the glaring signal that something may have gone wrong in the Michigan State Police crime lab in early 2002 (namely, the disconcerting Ruelas match). Nevertheless, we next assume the opposite. That is, that the lab procedures in early 2002 were so impeccable that only 1 in 15,000 analyses resulted in a cross-contamination event. Setting the

contamination rate to 1-in-15,000, the posterior probability of H_1 and H_2 come to 0.23 and 0.77, respectively. In other words, even then, the evidence points decidedly in favor of innocence.

What happens if we now bias the analysis in favor of an outcome of guilt to a seemingly absurd degree by assuming both that the number of plausible suspects prior to knowing who matched was only 250,000 (i.e., $N = 250{,}000$) *and* that the undetected cross-contamination rate in early 2002 was an impressive 1-in-15,000 analyses (despite the inexplicable match to Ruelas)? Even then, the probability of guilt is only 0.54 (i.e., probability of innocence = 0.46). In other words, to tip the scales ever-so-slightly in favor of guilt, we need to make multiple absurd assumptions designed to yield an outcome in favor of guilt.

The analysis summarized in Table 1 illustrates the essence of our Bayesian statistical analysis, but it is incomplete in one important respect. Specifically, the inputs to the analysis consisted of fixed estimates. For example, $P(E_s|H_1)$ – the probability that the DNA found on Mixer's pantyhose would be consistent with saliva – was fixed at 0.50. Although the estimates used in our analysis were all based on prior knowledge, their exact values are unknown, which raises a question: How would our conclusions change if we introduced variability into those inputs to reflect uncertainty as to their true values? The steps involved in taking into account uncertainty in the inputs are fairly technical and need not be considered by readers who are mainly interested in the take-home message. However, the analysis we have performed to this point would be incomplete if we did not investigate the degree to which our main conclusion is dependent on the exact values for the inputs, so we turn now to a consideration of that issue.

Modeling Uncertainty

To model uncertainty in the inputs of our Bayesian analysis, we created probability distributions for our estimates of $P(H_1)$, $P(E_m|H_1)$, $P(E_s|H_1)$, $P(E_e|H_1)$, $P(E_c|H_1)$, $P(H_2)$, $P(E_m|H_2)$,

and $P(E_e|H_2)$. The only estimates for which no variability was introduced are the ones that, according to H_2, must equal 1. More specifically, if H_2 is true, then the DNA on Mixer's clothing had to be consistent with saliva, so $P(E_s|H_2)$ remained fixed at 1.0. Similarly, according to H_2, the samples had to be contemporaneously analyzed in the lab, so $P(E_c|H_2)$ remained fixed at 1.0. We introduced variability in all other inputs by treating them as a distribution rather than as a constant.

Figure 1 shows the distributions we used for each input in this analysis. The top row shows the distributions for the inputs associated with H_1, with each panel corresponding to one of the 5 inputs, $P(H_1)$, $P(E_m|H_1)$, $P(E_s|H_1)$, $P(E_e|H_1)$, and $P(E_c|H_1)$. The bottom row shows the corresponding distributions for the inputs associated with H_2. The distributions for $P(E_s|H_2)$ and $P(E_c|H_2)$ are missing in the bottom row because those values were fixed at 1, as required by H_2. For each distribution, the black vertical line corresponds to the fixed value that was used previously in our main analysis (i.e., the black vertical lines correspond to the values shown in Table 1). For example, the top left panel of Figure 1 shows the distribution of the value of N we used that corresponds to our uncertainty in $P(H1)$. For our main analysis, we fixed N at 1 million, which is the value represented by the black vertical line. Other uncertainties are expressed directly as distributions on probabilities or distributions on rates, as in the cross-contamination rate. Our guiding principle in setting these variabilities was to be liberal, that is, to try to err on the side of too much variability. The ranges subtended seem to us to be the maximal plausible ranges.

The resultant probability of guilt is found by the mathematical operation of integrating the previous equations with respect to these distributions. It is convenient to perform this integration numerically, and we do so with a technique called Monte Carlo simulation. Here is

how it works. On each iteration, a value is sampled from each of the distributions in Figure 1. Then with these values, we used the equations shown Table 1 to compute the posterior odds of guilt for that iteration. We continued this process for 100,000 iterations. Because the set of inputs differ for each iteration, a different estimate of the posterior probability of guilt was produced each time, and the total resultant is a posterior distribution of the probability of guilt. Figure 2 shows this posterior, and it summarizes the results of this analysis. Note that despite considerable variability in the inputs (reflecting uncertainty as to their true values), the posterior probability of guilt ranges from ~0 to ~0.20 (i.e., estimated probability of innocence ranges from ~1.0 to ~0.80). Both the mean and the mode of this posterior distribution come to ~0.03. Thus, even after allowing for considerable uncertainty in the inputs to our analysis, the take-home message remains the same: the DNA evidence in this case point strongly in the direction of innocence.

Other Evidence

The preceding analysis considered only DNA evidence. We did so because our reading of the case suggests that the DNA match was decisive (e.g., the prosecutor's extensive closing arguments were almost entirely focused on the DNA evidence) and because it allowed us to illustrate how even a small rate of contamination can reverse the conclusion of a DNA match. However, once concerns about the DNA analysis are appreciated, it seems natural to wonder about other evidence – and other possible suspects – related to this case. We highlight here how additional considerations may be incorporated into our analyses.

Before the DNA match in the Michigan State Police Laboratory implicated Gary Leiterman, Jane Mixer was widely believed to have been murdered by John Norman Collins.[25] He was, after all, a serial killer known to have murdered several young women in the Ann Arbor

[25] *See supra* note 21.

area beginning in the summer of 1967 and ending in the summer of 1969 (Mixer was murdered in March of 1969). Figure 3 shows where the victims of the 7 murders originally attributed to Collins were found, and there is nothing in that map to suggest that Mixer was killed by someone else. Even though the prior on Collins is clearly higher than that of the other adult males who lived in the area in 1969, we did not take that fact into consideration in the analysis performed thus far. Instead, thus far, we have assumed that, for any randomly selected adult male living in the vicinity of the murder in 1969, including Collins, the prior probability that he was at the crime scene (and, by assumption, murdered Jane Mixer) was one in a million, or 10^{-6}.

In truth, the prior probability was not evenly distributed over the relevant population because Collins was widely believed to have murdered Jane Mixer before the DNA test results became known. Therefore, the prior probability that Collins was at the crime scene was considerably higher than 10^{-6}. That being the case, the prior probability associated with everyone else in the relevant population, including Leiterman, was considerably lower than 10^{-6}. Let p represent the prior probability that Collins was at the crime scene in 1969. The prior probability for all remaining $(10^6 - 1)$ males would now become $(1 - p) \times 10^{-6} / (1 - 10^{-6})$. However, if $p = .50$, then the prior for everyone else, including Leiterman, becomes $(1 - .50) \times 10^{-6} / (1 - 10^{-6}) \approx .50 \times 10^{-6}$ (about one in 2 million rather than one in 1 million). In other words, $p(H_1)$ in Table 1 would change from its current setting of 10^{-6} to $.50 \times 10^{-6}$.

With the prior probability for Collins set to p and the prior probabilities for the remaining males in the relevant population adjusted to $(1 - p) \times 10^{-6} / (1 - 10^{-6})$, the posterior odds that Leiterman deposited his DNA at the crime scene in 1969 would change from .02935 (the value shown in Table 1) to $.02935 \times (1 - p) / (1 - 10^{-6})$, and the posterior *probability* would change to $[.02935 \times (1 - p) / (1 - 10^{-6})] / [.02935 \times (1 - p) / (1 - 10^{-6}) + 1]$. Again using $p = .5$ as an

example, the posterior probability that Leiterman deposited his DNA at the crime scene would equal *P*(Guilt) = .0145, about half of the value we estimated previously (.0285 in Table 1).

We know of no empirical evidence that could be brought to bear on the estimate of *p*, which is why we performed our analysis with *p* set to 10^{-6}. However, setting it to a value like .50 does not seem entirely unreasonable. It would mean that, if it happened to be available, the relevant empirical evidence would show that when a murder victim who fits the profile of the other victims of a serial killer is found in the middle of that serial killer's murder spree and in the same general vicinity as the other victims of that serial killer (see Figure 3), then, about half the time, that victim, too, will turn out to have been murdered by the serial killer. Whatever the actual value of *p* might be, it seems certain that it is much higher than 10^{-6}. By setting it to a more realistic value, we can not only update our posterior estimate of *P*(Guilt), as we did above, but we can also obtain some idea of the relative posterior probability of Collins vs. Leiterman being the one who murdered Jane Mixer.

Computing the relative posterior probability for Collins vs. Leiterman requires not only that the posterior probability for Leiterman be computed; it also requires that the posterior probability for Collins be computed (i.e., the probability that Collins was at the crime scene despite the DNA match to Leiterman). Note that for any setting of *p* equal to or greater than 10^{-6}, the posterior probability of guilt for Leiterman, while always small, will still be much larger than his prior probability. For example, with $p = 10^{-6}$ for Collins (its original value), the prior probability for Leiterman, $p(H_1)$ is also 10^{-6}, whereas the posterior probability for Leiterman (as shown in Table 1) increases to *P*(Guilt) = 0.0285. This posterior probability is small (which is the main point of our article) but is also much larger than the prior probability of 10^{-6}. Thus, while the probability that he left his DNA at the crime scene is slight, it is clearly higher than it

was before the match. Similarly, as described above, if $p = .5$ (an arguably more reasonable prior for Collins), the prior probability for Leiterman, $p(H_1)$, becomes $(1 - .5) \times 10^{-6} / (1 - 10^{-6}) \approx .5 \times 10^{-6}$, and the posterior probability, P(Guilt), becomes .0145. Again, despite being small, this posterior probability is still much larger than the prior probability. Because the posterior probability for Leiterman is much larger than his prior probability for any reasonable setting of p, the posterior probability for the rest of the relevant population, including the posterior probability for Collins, would now have to decrease slightly so that the posterior probabilities over the entire relevant population would sum to 1. The posterior probability for everyone other than Leiterman is obtained by multiplying their prior probability by $1 - [P(\text{Guilt}) - P(H_1)] / [1 - P(H_1)]$, where both P(Guilt) and $P(H_1)$ are values that correspond to Leiterman and that depend on p. Thus, if we again set $p = .5$ such that P(Guilt) = .0145 and $P(H_1) \approx .50 \times 10^{-6}$, this equation indicates that the prior probabilities for everyone else would be multiplied by $1 - (.0145 - .50 \times 10^{-6}) / (1 - .50 \times 10^{-6}) = .9855$. For Collins, his prior probability of $p = .50$ would drop to a posterior probability of $.50 \times .9855 = .4928$ as a result of the Leiterman DNA match.

We can now use these values to compute the posterior odds that Collins vs. Leiterman was at the crime scene (and, by assumption, murdered Jane Mixer) in 1969. With $p = .50$, the posterior odds that Collins vs. Leiterman was at the crime scene in 1969 would be $.4928 / .0145 = 33.99$. In other words, Collins was ~34 times more likely than Leiterman to have been at the crime scene, despite the DNA match to Leiterman. With $p = .10$ (which seems like an implausibly low setting given that Collins alone was strongly suspected of having murdered Jane Mixer before the DNA test results became known), the posterior odds that Collins vs. Leiterman was at the crime scene in 1969 would be $.0974 / .0257 = 3.79$ (i.e., Collins would still be ~4 times more likely than Leiterman to have been at the crime scene). The value of p would have to

be set as low as .0285 for the odds to become even.

No matter what the estimate of *p*, one advantage of explicitly considering the role of Collins is that it allows us to address the implications of what, on the surface, appears to be a somewhat incriminating piece of non-DNA evidence against Leiterman, namely, the fact that in 1969, he owned a .22 caliber gun (the kind of gun that was used to kill Jane Mixer). However, John Norman Collins also owned a .22 caliber gun at the time, one that he used to kill another University of Michigan graduate student (by shooting her in the head) just 10 weeks after someone killed University of Michigan graduate student Jane Mixer (also by shooting her in the head with a .22). Thus, ownership of a .22 caliber gun in 1969, while constituting some evidence against both Leiterman and Collins, does not tip the scales in the direction of either one as the person who murdered Jane Mixer. Instead, this evidence variable simply cancels out.

Conclusion

The bottom line of our statistical inquiry into this matter is that if we assume a contamination rate in the Michigan State Police Laboratory in 2002 of 0 despite the disconcerting match to John Ruelas, then the outcome of the analysis suggests that Gary Leierman deposited his DNA on the Mixer evidence in 1969 (which in turn would mean that he very likely murdered her). However, assuming a contamination rate as low as 1-in-1500 (99.93% of analyses are error-free), then the outcome of the analysis suggests that Gary Leierman's DNA was deposited on the Mixer evidence through contamination in 2002 with probability .97 (which in turn would mean that John Norman Collins very likely murdered her). The complete reversal of the implication of a contamination rate of 0 and a very small contamination rate is utterly non-intuitive and may help to explain why Leiterman was convicted by a jury.

The analysis summarized in Table 1 is dependent on our reading of the empirical

evidence used to estimate the various parameters upon which the analysis was based. If our reading of that evidence turns out to be incorrect, or if new findings come to light showing that the evidence we relied upon is not valid, then the result of our analysis (summarized in Table 1 and Figure 2) would change accordingly. In addition, our analysis is dependent on the assumption that the pertinent events are the four events listed in Table 1. If other events are deemed to be relevant, then they would need to be taken into consideration as well, potentially changing the outcome of the analysis. However, these considerations do not change the main point of our analysis, which is that under the right conditions, the difference between no chance of contamination and a small chance of contamination can be the difference between no chance of innocence and almost certain innocence.

It is important to emphasize that our analysis does not in any way call DNA evidence into doubt as a general rule. The circumstances of the Leiterman case were special, though perhaps not unique. Had the two DNA analyses been performed in different labs, or at very different times, as would often be the case, the contamination rate could well be infinitesimal. Likewise, if the DNA was from a sample incompatible with the reference sample (saliva), then, again, there would be no concern. Furthermore, if the various details of the relative balance between the victim's DNA on the evidence and the suspect's DNA on the evidence had been different from what is was in Leiterman's case (e.g., if much more of the wearer's than the toucher's DNA had been observed), it would point less to innocence. Finally, if there had been, as in this case there was not, some convincing evidence discovered after the DNA match made someone a compelling suspect, then the case could be, even with some non-zero chance of an initial contamination, sufficient for a guilty verdict. But in this case, intuition notwithstanding, the relevant data align and clearly point in the direction of contamination and, it would therefore

seem, in the direction of innocence.

Table 1. *A summary of the Bayesian analysis of the Gary Leiterman/Jane Mixer case*.

Hypotheses:

- H_1: Leiterman's DNA was deposited on Mixer's clothing only at the crime scene in 1969
- H_2: Leiterman's DNA was deposited on Mixer's clothing only in the lab in 2002
- H_3: Leiterman's DNA was deposited on Mixer's clothing at the crime scene and in the lab

Events:

- E_m: A match between Leiterman's DNA profile and the forensic DNA profile.
- E_s: The forensic DNA sample was consistent with saliva.
- E_e: Leiterman's DNA was exclusively found on Mixer's clothing.
- E_c: Leiterman's reference DNA sample and Mixer's forensic evidence contemporaneously analyzed.

Prior Odds (H_1 vs. H_2):

- $P(H_1) = 1 / 1 \text{ million} = 10^{-6}$
- $P(H_2) = 1 / 5000 = 2 \times 10^{-4}$
- $P(H_1)/ P(H_2) = 0.005$ (i.e., 1 / 200)

Likelihood Ratio:

- $P(E_m|H_1) = 0.25$ (Leiterman's DNA left at scene and survives 33 years, per H_1)
- $P(E_m|H_2) = 1 / 1500 = 6.67 \times 10{-}4$ (cross-contamination rate in the lab, per H_2)
- $P(E_s|H_1) = 0.50$
- $P(E_s| H_2) = 1.0$
- $P(E_e|H_1) = 0.15$
- $P(E_e|H_2) = 0.575$
- $P(E_c|H_1) = 0.12$
- $P(E_c| H_2) = 1$
- $P(D|H_1) = P(E_m|H_1) \times P(E_s|H_1) \times P(E_e|H_1) \times P(E_c|H_1) = 0.00225.$
- $P(D|H_2) = P(E_m|H_2) \times P(E_s|H_2) \times P(E_e|H_2) \times P(E_c|H_2) = 0.000383$
- $P(D|H_1) / P(D|H_2) = 5.87$

Posterior Odds of Guilt:

- $P(H_1|D)/ P(H_2|D) = P(H_1)/ P(H_2) \times P(D|H_1) / P(D|H_2) = 0.0294$

Posterior Probability of Guilt:

- $P(H_3|D) = P(H_1|D) \times P(H_2|D) \approx 0$
- $P(\text{Guilt}) = [P(H_1|D) + P(H_3|D)] / [P(H_1|D) + P(H_2|D) + P(H_3|D)]$
- $P(\text{Guilt}) \approx P(H_1|D) / [P(H_1|D) + P(H_2|D)] = 0.0294 / (0.0294 + 1) = 0.0285$

Figure 1. Distributions of inputs used to compute the posterior distribution of the probability of guilt. The median of each distribution is denoted by a black vertical line, which also corresponds to the fixed values used for the Bayesian analysis summarized in Table 1.

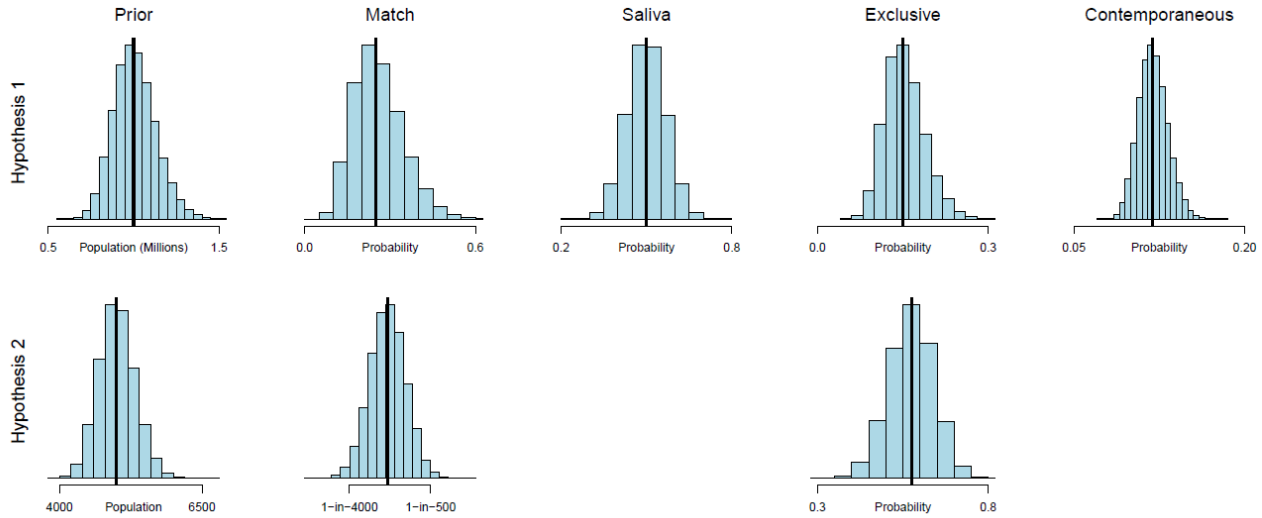


Figure 2. Posterior distribution of the probability of guilt for a Bayesian analysis performed using the distributions depicted in Figure 1.

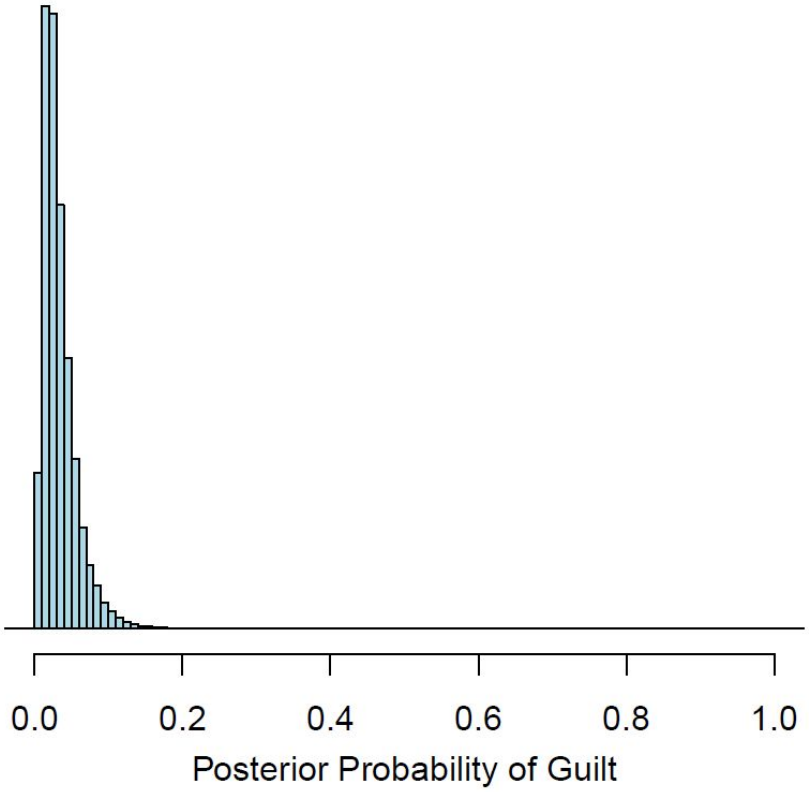
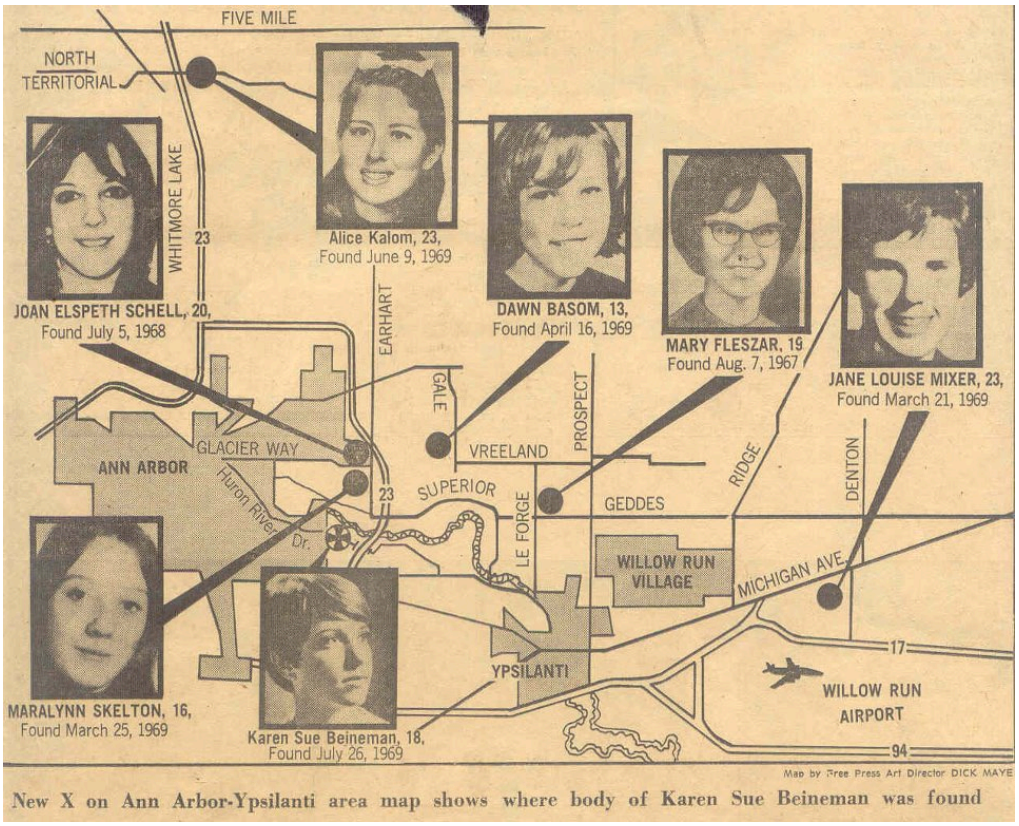


Figure 3. A map showing where the suspected victims of John Norman Collins (and Jane Mixer) were found. Alice Kalom and Jane Mixer were both University of Michigan graduate students, and both were shot in the head with a .22. Mixer was the 3rd of the 7 murder victims shown on this map. John Norman Collins was convicted of murdering the last of the 7, Karen Sue Beineman.



New X on Ann Arbor-Ypsilanti area map shows where body of Karen Sue Beineman was found

Appendix A

An alternative approach to estimating the value of *N* can be obtained use DNA database statistics. This estimate of *N* will represent the size of the active criminal population in Michigan in 1969, a subset of which had their DNA profiles in the database when Leiterman's profile matched in 2004. The Michigan DNA database today, which is part of CODIS, currently contains 373,874 known offender profiles plus 51,573 known arrestee profiles for a total of 373,874 + 51,573 = 425,447 known profiles. When an unknown profile from crime scene evidence is entered into this database, and a match is obtained, an investigation is aided. When no match occurs, the still unknown profile is added to the CODIS database as a non-identified "forensic profile." Currently, in Michigan, 11,792 investigations have been aided from matches to known profiles in the database. In addition, there are 25,427 unknown forensic profiles in the database. Thus, of the 11,792 + 25,427 = 37,219 searches, 11,792 / 37219 = .32 matched a known profile (the kind of match that happened when the unknown profile from the Mixer evidence matched the known profile in the database of Gary Leiterman).

If every member of the active criminal population in Michigan today were already in the database, then an unknown forensic profile, when recovered from a crime scene and entered into the database, would always match the known profile of the perpetrator. In that case, it would follow that *N* (the size of the active criminal population in Michigan before knowing who matched) would equal the current number of profiles in the database, or *N* = 425,447. However, a match occurs only about 1/3 of the time. This result suggests that a better estimate of the active criminal population in Michigan today is about 3 times the number of individuals

already in the database, or 3 × 425,447 = 1,276,341 individuals. As noted above, the population of Michigan in 1969 was approximately 89% of what it is today, so a reasonable estimate of the active criminal population in Michigan back in 1969 is N = .89 × 1,276,341 = 1,135,944.

This estimate of N uses the highly simplified Lincoln-Peterson mark-recapture formula that has long been used to estimate the size of uncountable wildlife populations, such as the number of fish in a lake.[26] The estimate is only approximate, but it has the virtue of being objective, and the resulting value seems intuitively reasonable. A more exact estimate would require a model that takes into account the fact that members of active criminal population die, new criminals come of age, different criminals have different probabilities of being captured, etc. However, this estimate corresponds closely to the estimate we derived based on geographical considerations (N = 1 million), so it seems like a reasonable figure to use for our main analysis.

[26] Byron K. Williams, J. D. Nichols & M. J. Conroy, *Analysis and management of animal populations.* Academic Press: San Diego (2001).